

AD-A170 256

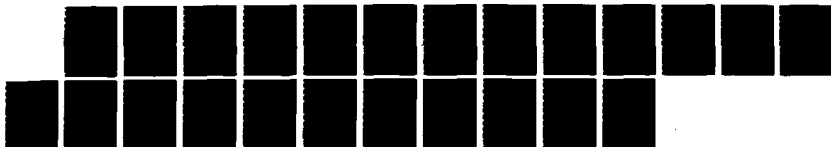
CONVERGENCE THEOREMS OF SAMPLER PROCESSES WITH  
APPLICATIONS TO IMAGE PROCESSING(U) NORTH CAROLINA UNIV  
AT CHAPEL HILL DEPT OF STATISTICS K VANAGI JAN 86  
N00014-81-K-0173

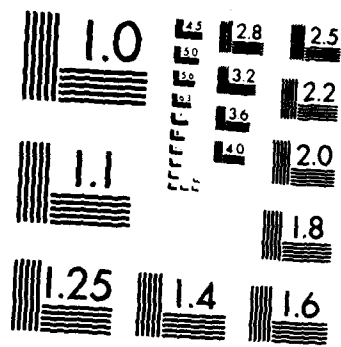
1/1

UNCLASSIFIED

F/G 12/1

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A170 256

DTIC FILE COPY

12

Convergence Theorems of Sampler Processes  
with Applications to Image Processing (\*)

Kenjiro Yanagi

Department of Mathematics  
Faculty of Science  
Yamaguchi University  
Yamaguchi 753  
Japan

LISS 15  
January, 1986

DTIC  
ELECTE  
JUL 28 1986  
S B D

DISTRIBUTION STATEMENT A  
Approved for public release  
Distribution Unlimited

(\*)These results are based on research done while the author was visiting at the Department of Statistics, University of North Carolina at Chapel Hill as an overseas researcher of the Japanese Ministry of Education

This research was partially supported by ONR Contracts N00014-81-K-0373 and N00014-84-C-0212.

86 7 28 163

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A170256

## REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION <b>UNCLASSIFIED</b>			1b. RESTRICTIVE MARKINGS										
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION/AVAILABILITY OF REPORT  Approved for Public Release: Distribution Unlimited										
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			5. MONITORING ORGANIZATION REPORT NUMBER(S)										
4. PERFORMING ORGANIZATION REPORT NUMBER(S)			7a. NAME OF MONITORING ORGANIZATION										
6a. NAME OF PERFORMING ORGANIZATION  Department of Statistics		6b. OFFICE SYMBOL (If applicable)		7b. ADDRESS (City, State and ZIP Code)									
6c. ADDRESS (City, State and ZIP Code)  University of North Carolina Chapel Hill, North Carolina 27514		8a. NAME OF FUNDING/SPONSORING ORGANIZATION  Office of Naval Research		8b. OFFICE SYMBOL (If applicable)									
6c. ADDRESS (City, State and ZIP Code)  Statistics & Probability Program Arlington, VA 22217		9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER  N00014-81-K-0373 & N0014-84-C-0212											
11. TITLE (Include Security Classification) Convergence Theorems of Sampler Processes . . .		10. SOURCE OF FUNDING NOS.											
		<table border="1"> <tr> <th>PROGRAM ELEMENT NO.</th> <th>PROJECT NO.</th> <th>TASK NO.</th> <th>WORK UNIT NO.</th> </tr> <tr> <td>NR</td> <td>042</td> <td>269</td> <td>SRO 105</td> </tr> </table>				PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.	NR	042	269	SRO 105
PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.										
NR	042	269	SRO 105										
12. PERSONAL AUTHOR(S)  Kenjiro Yanagi													
13a. TYPE OF REPORT  TECHNICAL		13b. TIME COVERED FROM _____ TO _____		14. DATE OF REPORT (Yr., Mo., Day) Jan., 1986									
15. PAGE COUNT 19													
16. SUPPLEMENTARY NOTATION													
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)										
FIELD	GROUP	SUB. GR.	Gibbs distribution, Markov random field, image processing										
19. ABSTRACT (Continue on reverse if necessary and identify by block number)  We consider sampling methods for multidimensional random fields. We study the convergence of a sampler process generated by the methods stated in Introduction. We can apply convergence theorem to the restoration of degraded images in image processing.													
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT  UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input checked="" type="checkbox"/> OTIC USERS <input type="checkbox"/>			21. ABSTRACT SECURITY CLASSIFICATION										
22a. NAME OF RESPONSIBLE INDIVIDUAL  C.R. Baker			22b. TELEPHONE NUMBER (Include Area Code) (919) 962-2189		22c. OFFICE SYMBOL								

**Abstract.**

We consider sampling methods for multidimensional random fields. We study the convergence of a sampler process generated by the methods stated in Introduction. We can apply convergence theorem to the restoration of degraded images in image processing.

**Key words.** Gibbs distribution, Markov random field,  
image processing

Accession Date	
No. of Pages	✓
Author	
Title	
Subject	
Abstract	
Notes	
D.L. 1	
A-1	



## 1. Introduction.

We imagine a simple processor placed at each site  $s$  of the graph. The state of the machine evolves by discrete changes and it is therefore convenient to discretize time, say  $t = 1, 2, 3, \dots$ . At time  $t$ , the state of the processor at site  $s$  is a random variable  $X_s(t)$  with values in  $\Lambda_s = \Lambda = \{0, 1, 2, \dots, L-1\}$ . The total configuration is  $X(t) = (X_{s_1}(t), X_{s_2}(t), \dots, X_{s_N}(t))$ , which evolves due to state changes of the individual processors. The starting configuration,  $X(0)$ , is arbitrary. At each epoch, only one site undergoes a change, so that  $X(t-1)$  and  $X(t)$  can differ in at most one coordinate. Let  $n_1, n_2, \dots$  be the sequence in which the sites are visited for replacement: thus  $n_t \in S$  and  $X_{s_i}(t) = X_{s_i}(t-1)$ ,  $i \neq n_t$ . Each processor is programmed to follow the same algorithm: at time  $t$ , a sample is drawn from the local characteristics of not necessarily Gibbs measure  $\pi_t$  for  $s = n_t$  and  $\omega = X(t-1)$ . In other words, we choose a state  $x \in \Lambda_{n_t}$  from the conditional distribution of  $X_{n_t}$  given the observed states of the sites  $X_r(t-1)$ ,  $r \neq n_t$ . The new configuration  $X(t)$  has  $X_{n_t}(t) = x$  and  $X_s(t) = X_s(t-1)$ ,  $s \neq n_t$ . Given an initial configuration,  $X(0)$ , we obtain a sequence  $X(1), X(2), \dots$  of configurations which converge to a limit distribution  $\pi_\infty$  of  $\pi_t$ . The limits obtained do not depend on  $X(0)$ . When  $\pi_t$  is Gibbs measure, we can apply our theorems to the annealing restorations of degraded images.

The author wishes to express his gratitude to Prof. H. Umegaki and Prof. C. R. Baker for interest in and valuable comments to the present work.

## 2. Preliminaries.

Let  $S = \{s_1, s_2, \dots, s_N\}$  be a set of sites and let  $G = \{G_s, s \in S\}$  be a neighborhood system for  $S$ , meaning any collection of subsets of  $S$  for which 1)  $s \notin G_s$  and 2)  $s \in G_r \Leftrightarrow r \in G_s$ . Obviously,  $G_s$  is the set of neighbors of  $s$  and the pair  $\{S, G\}$  is a graph in the usual way. A subset  $C \subset S$  is a clique if every pair of distinct sites in  $C$  are neighbors;  $C$  denotes the set of cliques. Let  $X = \{X_s, s \in S\}$  denote any family of random variables indexed by  $S$ . For simplicity, we can assume a common state space, say  $\Lambda = \{0, 1, 2, \dots, L-1\}$ , so that  $X_s \in \Lambda$  for all  $s$ . Let  $\Omega$  be the set of all possible configurations:

$$\Omega = \{\omega = (x_{s_1}, \dots, x_{s_N}) : x_{s_i} \in \Lambda, 1 \leq i \leq N\}.$$

As usual, the event  $\{X_{s_1} = x_{s_1}, \dots, X_{s_N} = x_{s_N}\}$  is abbreviated  $\{X = \omega\}$ .

$X$  is a Markov random field (we abbreviate MRF) with respect to  $G$  if

$$P(X = \omega) > 0 \text{ for all } \omega \in \Omega; \quad (2.1)$$

$$P(X_s = x_s | X_r = x_r, r \neq s) = P(X_s = x_s | X_r = x_r, r \in G_s) \quad (2.2)$$

for every  $s \in S$  and  $(x_{s_1}, \dots, x_{s_N}) \in \Omega$ . Technically, what is meant here is that the pair  $\{X, P\}$  satisfies (2.1) and (2.2) relative to some probability measure on  $\Omega$ . The collection of functions on the left-hand side of (2.2) is called the local characteristics of the MRF and it turns out that the joint probability distribution  $P(X = \omega)$  of any process satisfying (2.1) is uniquely determined by these conditional probabilities.

A Gibbs distribution with respect to  $\{S, G\}$  is a probability measure  $\pi$  on  $\Omega$  with the following representation:

$$\pi(\omega) = \frac{1}{Z} \exp(-U(\omega)/T)$$

where  $Z$  and  $T$  are constants and  $U$ , called the energy function, is of the form

$$U(\omega) = \sum_{C \in \mathcal{C}} V_C(\omega).$$

Each  $V_C$  is a function on  $\Omega$  with the property that  $V_C(\omega)$  depends only on those coordinates  $x_s$  of  $\omega$  for which  $s \in C$ . Such a family  $\{V_C, C \in \mathcal{C}\}$  is called a potential.  $Z$  is the normalizing constant:

$$Z = \sum_{\omega} \exp(-U(\omega)/T)$$

and is called the partition function.  $T$  stands for "temperature"; for our purposes,  $T$  controls the degree of "peaking" in the "density"  $\pi$ . The following proposition gives the equivalence between MRF and Gibbs distribution. For a proof see [1],[6].

PROPOSITION 1. Let  $G$  be a neighborhood system. Then  $X$  is an MRF with respect to  $G$  if and only if  $\pi(\omega) = P(X = \omega)$  is a Gibbs distribution with respect to  $G$ .

We define a sample process  $\{X(t); t = 0, 1, 2, \dots\}$  in the following: Let  $n_1, n_2, \dots$  be the sequence in which the sites are visited for updating. The initial configuration is  $X(0)$ . The evolution  $X(t-1) \rightarrow X(t)$  of the system is defined by

- (1)  $X(t-1)$  and  $X(t)$  can differ in at most one coordinate  $n_t$ ;
- (2)  $\{X(t); t = 0, 1, 2, \dots\}$  is generated by

$$P(X_s(t)=x_s, s \in S) = \pi_t(x_{n_t} | x_s, s \neq n_t) P(X_s(t-1)=x_s, s \neq n_t)$$

where  $\pi_t$  is a probability measure on  $\Omega$  such that  $0 < \pi_t(\omega) < 1$  for every  $\omega \in \Omega$ .

Let  $\|\mu - \nu\|$  denote the  $L^1$ -distance between two distributions on  $\Omega$ :

$$\|\mu - \nu\| = \sum_{\omega} |\mu(\omega) - \nu(\omega)|.$$

Obviously  $\mu_n \rightarrow \mu$  ( $n \rightarrow +\infty$ ) in distribution (i.e.,  $\mu_n(\omega) \rightarrow \mu(\omega)$  for every  $\omega$ ) if and only if  $\|\mu_n - \mu\| \rightarrow 0$  ( $n \rightarrow +\infty$ ).

We assume that  $\pi_t$  has a limit  $\pi_\infty$  as  $t \rightarrow +\infty$ .

### 3. Convergence Theorem.

In this section we state the main theorem and give the proof.

THEOREM 1. Assume that there exists an integer  $\tau \geq N$  such that for every  $t = 0, 1, 2, \dots$  we have

$$S \subset \{n_{t+1}, n_{t+2}, \dots, n_{t+\tau}\}.$$

Suppose that

$$\delta(t) = \inf\{\pi_t(x_{s_i} | x_{s_j}, j \neq i); 1 \leq i \leq N, (x_{s_1}, \dots, x_{s_N}) \in \Omega\}$$

$$\geq \frac{C}{N\sqrt{t}}$$

for some constant  $C > 0$ .

Furthermore, suppose that  $\pi_t(\omega)$  is monotone for all  $t \geq t'$  for some integer  $t'$  and for all  $\omega \in \Omega$ . Then for any starting configuration  $\eta \in \Omega$  and for every  $\omega \in \Omega$ ,

$$\lim_{t \rightarrow +\infty} P(X(t)=\omega | X(0)=\eta) = \pi_\infty(\omega).$$

We need the following two lemmas to prove the above theorem.

LEMMA 1. For every  $t_0 = 0, 1, 2, \dots$ ,

$$\lim_{t \rightarrow +\infty} \sup_{\omega, \eta', \eta''} |P(X(t)=\omega | X(t_0)=\eta') - P(X(t)=\omega | X(t_0)=\eta'')| = 0$$

PROOF. Fix  $t_0 = 0, 1, 2, \dots$  and define  $T_k = t_0 + k\tau$ ,  $k = 0, 1, 2, \dots$ . By hypothesis  $S \subset \{n_{t+1}, \dots, n_{t+\tau}\}$  for all  $t$ . Now fix  $k$  for the moment and define the  $m_i$  as follows:

$$m_i = \sup\{t; t \leq T_k, n_t = s_i\}, 1 \leq i \leq N.$$

We assume that  $m_1 > m_2 > \dots > m_N$ . Then

$$\begin{aligned}
& P(X(T_k) = \omega | X(T_{k-1}) = \omega') \\
&= P(X_{s_1}(T_k) = x_{s_1}, \dots, X_{s_N}(T_k) = x_{s_N} | X(T_{k-1}) = \omega') \\
&= P(X_{s_1}(m_1) = x_{s_1}, \dots, X_{s_N}(m_N) = x_{s_N} | X(T_{k-1}) = \omega') \\
&= \prod_{j=1}^N P(X_{s_j}(m_j) = x_{s_j} | \\
&\quad X_{s_{j+1}}(m_{j+1}) = x_{s_{j+1}}, \dots, X_{s_N}(m_N) = x_{s_N}, X(T_{k-1}) = \omega') \\
&\geq \prod_{j=1}^N \delta(m_j) \geq C^N(t_0 + k\tau)^{-1}.
\end{aligned}$$

We obtain

$$\inf_{\omega, \omega'} P(X(T_k) = \omega | X(T_{k-1}) = \omega') \geq C^N(t_0 + k\tau)^{-1} \quad (3.1)$$

for every  $t_0 = 0, 1, 2, \dots$  and  $k = 1, 2, \dots$ , bearing in mind that

$T_k$  depends on  $t_0$ . For each  $t > 0$ , we define  $K(t) = \sup\{k;$

$T_k < t\}$  so that  $K(t) \rightarrow +\infty$  as  $t \rightarrow +\infty$ .

For fixed  $t > T_1$ ,

$$\begin{aligned}
& \sup_{\omega, \eta', \eta''} |P(X(t) = \omega | X(0) = \eta') - P(X(t) = \omega | X(0) = \eta'')| \\
&= \sup_{\omega} \{ \sup_{\eta} P(X(t) = \omega | X(0) = \eta) - \inf_{\eta} P(X(t) = \omega | X(0) = \eta) \} \\
&= \sup_{\omega} \{ \sup_{\eta} \sum_{\omega'} P(X(t) = \omega | X(T_1) = \omega') P(X(T_1) = \omega' | X(0) = \eta) \\
&\quad - \inf_{\eta} \sum_{\omega'} P(X(t) = \omega | X(T_1) = \omega') P(X(T_1) = \omega' | X(0) = \eta) \} \\
&= \sup_{\omega} \Omega(t, \omega).
\end{aligned}$$

Certainly, for each  $\omega \in \Omega$ ,

$$\begin{aligned}
& \sup_{\eta} \sum_{\omega'} P(X(t) = \omega | X(T_1) = \omega') P(X(T_1) = \omega' | X(0) = \eta) \\
&\leq \sup_{\mu} \sum_{\omega'} P(X(t) = \omega | X(T_1) = \omega') \mu(\omega')
\end{aligned}$$

where the supremum is over all probability measure  $\mu$  on  $\Omega$  which, by (3.1), are subject to  $\mu(\omega') \geq C^N(t_0 + k\tau)^{-1}$  for every  $\omega' \in \Omega$ . Suppose  $\omega' \rightarrow P(X(t)=\omega | X(T_1)=\omega')$  is maximized at  $\omega' = \omega^*$  (which depends on  $\omega$ ). Then the last supremum is attained by

$$\begin{aligned}\mu(\omega^*) &= 1 - (L^N - 1)C^N(t_0 + k\tau)^{-1} \\ \mu(\omega') &= C^N(t_0 + k\tau)^{-1}, \quad \omega' \neq \omega^*.\end{aligned}$$

The value so obtained is

$$\begin{aligned}&(1 - (L^N - 1)C^N(t_0 + k\tau)^{-1})P(X(t)=\omega | X(T_1)=\omega^*) \\ &+ C^N(t_0 + k\tau)^{-1} \sum_{\omega' \neq \omega^*} P(X(t)=\omega | X(T_1)=\omega').\end{aligned}$$

Similarly,

$$\begin{aligned}&\inf_{\eta} \sum_{\omega'} P(X(t)=\omega | X(T_1)=\omega') P(X(T_1)=\omega' | X(0)=\eta) \\ &\geq \inf_{\mu} \sum_{\omega'} P(X(t)=\omega | X(T_1)=\omega') \mu(\omega')\end{aligned}$$

where the infimum is over all probability measure  $\mu$  on  $\Omega$  which, by (3.1), are subject to  $\mu(\omega') \geq C^N(t_0 + k\tau)^{-1}$  for every  $\omega' \in \Omega$ . Suppose  $\omega' \rightarrow P(X(t)=\omega | X(T_1)=\omega')$  is minimized at  $\omega' = \omega_*$  (which depends on  $\omega$ ). Then the last infimum is attained by

$$\begin{aligned}\mu(\omega_*) &= 1 - (L^N - 1)C^N(t_0 + k\tau)^{-1} \\ \mu(\omega') &= C^N(t_0 + k\tau)^{-1}, \quad \omega' \neq \omega_*.\end{aligned}$$

The value so obtained is

$$\begin{aligned}&(1 - (L^N - 1)C^N(t_0 + k\tau)^{-1})P(X(t)=\omega | X(T_1)=\omega_*) \\ &+ C^N(t_0 + k\tau)^{-1} \sum_{\omega' \neq \omega_*} P(X(t)=\omega | X(T_1)=\omega').\end{aligned}$$

It follows that

$$\Omega(t, \omega) \leq$$

$$(1 - (L^N - 1)C^N(t_0 + k\tau)^{-1}) \{P(X(t)=\omega | X(T_1)=\omega^*) - P(X(t)=\omega | X(T_1)=\omega_*)\}$$

$$\begin{aligned}
& + C^N(t_0 + k\tau)^{-1} \left\{ \sum_{\omega' \neq \omega_*} P(X(t) = \omega | X(T_1) = \omega') - \sum_{\omega' \neq \omega_*} P(X(t) = \omega | X(T_1) = \omega') \right\} \\
& = (1 - L^N C^N(t_0 + k\tau)^{-1}) \{P(X(t) = \omega | X(T_1) = \omega_*) - P(X(t) = \omega | X(T_1) = \omega_*)\}.
\end{aligned}$$

Hence

$$\begin{aligned}
& \sup_{\omega, \eta', \eta''} |P(X(t) = \omega | X(0) = \eta') - P(X(t) = \omega | X(0) = \eta'')| \\
& \leq (1 - L^N C^N(t_0 + k\tau)^{-1}) \\
& \quad \sup_{\omega} \{P(X(t) = \omega | X(0) = \omega_*) - P(X(t) = \omega | X(0) = \omega_*)\} \\
& = (1 - L^N C^N(t_0 + k\tau)^{-1}) \\
& \quad \sup_{\omega, \eta', \eta''} |P(X(t) = \omega | X(T_1) = \eta') - P(X(t) = \omega | X(T_1) = \eta'')|.
\end{aligned}$$

Proceeding in this way, we obtain the bound

$$\prod_{k=1}^{K(t)} (1 - L^N C^N(t_0 + k\tau)^{-1}).$$

Hence it will be sufficient to show that

$$\lim_{m \rightarrow +\infty} \prod_{k=1}^m (1 - L^N C^N(t_0 + k\tau)^{-1}) = 0 \quad (3.2)$$

for every  $t_0$ . However (3.2) is a well-known consequence of the divergence of the series  $\sum_k (t_0 + k\tau)^{-1}$  for all  $t_0, t$ . This completes the proof of Lemma 1.

Q.E.D.

It will be convenient to use the following, semistandard notation for transitions. For nonnegative integers  $r < t$  and  $\omega, \eta \in \Omega$ , set

$$P(t, \omega | r, \eta) = P(X(t) = \omega | X(r) = \eta)$$

and, for any distribution  $\mu$  on  $\Omega$ , set

$$P(t, \omega | r, \mu) = \sum_{\eta} P(t, \omega | r, \eta) \mu(\eta).$$

LEMMA 2.  $\lim_{t_0 \rightarrow +\infty} \sup_{t \geq t_0} \|P(t, \cdot | t_0, \pi_\infty) - \pi_\infty\| = 0.$

PROOF. The probability measures  $P(t, \cdot | t_0, \pi_\infty)$  figure prominently in the proof, and for notational ease we prefer to write  $P_{t_0, t}(\cdot)$  such that for any  $t \geq t_0 \geq 0$  we have

$$P_{t_0, t}(\omega) = \sum_{\eta} P(X(t)=\omega | X(t_0)=\eta) \pi_\infty(\eta).$$

To begin with, we claim that for any  $t > t_0 \geq 0$ ,

$$\|P_{t_0, t} - \pi_t\| \leq \|P_{t_0, t-1} - \pi_t\| \quad (3.3)$$

Assume for convenience that  $n_t = s_1$ . Then

$$\begin{aligned} & \|P_{t_0, t} - \pi_t\| \\ &= \sum_{(x_{s_1} \dots x_{s_N})} |\pi_t(x_{s_1} | x_s, s \neq s_1) P_{t_0, t-1}(X_s = x_s, s \neq s_1) \\ & \quad - \pi_t(x_s, s \in S)| \\ &= \sum_{(x_{s_2} \dots x_{s_N})} \left\{ \sum_{x_{s_1} \in \Lambda} \pi_t(x_{s_1} | x_s, s \neq s_1) \right. \\ & \quad \left. |P_{t_0, t-1}(X_s = x_s, s \neq s_1) - \pi_t(x_s, s \neq s_1)| \right\} \\ &= \sum_{(x_{s_2} \dots x_{s_N})} |P_{t_0, t-1}(X_s = x_s, s \neq s_1) - \pi_t(x_s, s \neq s_1)| \\ &= \sum_{(x_{s_2} \dots x_{s_N})} \left| \sum_{x_{s_1}} \{ P_{t_0, t-1}(X_s = x_s, s \in S) - \pi_t(x_s, s \in S) \} \right| \\ &\leq \sum_{(x_{s_1} \dots x_{s_N})} |P_{t_0, t-1}(X_s = x_s, s \in S) - \pi_t(x_s, s \in S)| \\ &= \|P_{t_0, t-1} - \pi_t\|. \end{aligned}$$

Fix  $t > t_0 \geq 0$ ,

$$\begin{aligned} & \|P_{t_0, t} - \pi_\infty\| \\ &\leq \|P_{t_0, t} - \pi_t\| + \|\pi_t - \pi_\infty\| \end{aligned}$$

$$\leq \|P_{t_0, t-1} - \pi_t\| + \|\pi_t - \pi_\infty\| \quad \text{by (3.3)}$$

$$\leq \|P_{t_0, t-1} - \pi_{t-1}\| + \|\pi_{t-1} - \pi_t\| + \|\pi_t - \pi_\infty\|$$

$$\leq \|P_{t_0, t-2} - \pi_{t-1}\| + \|\pi_{t-1} - \pi_t\| + \|\pi_t - \pi_\infty\|$$

$$\leq \|P_{t_0, t-1} - \pi_{t-2}\| + \|\pi_{t-2} - \pi_{t-1}\| + \|\pi_{t-1} - \pi_t\| + \|\pi_t - \pi_\infty\|$$

Proceeding in this way,

$$\|P_{t_0, t} - \pi_\infty\| \leq \|P_{t_0, t_0} - \pi_{t_0}\| + \sum_{k=t_0}^{t-1} \|\pi_k - \pi_{k+1}\| + \|\pi_t - \pi_\infty\|$$

Since  $P_{t_0, t_0} = \pi_\infty$  and  $\|\pi_t - \pi_\infty\| \rightarrow 0$  as  $t \rightarrow +\infty$ , we have

$$\begin{aligned} & \overline{\lim}_{t_0 \rightarrow \infty} \sup_{t \geq t_0} \|P_{t_0, t} - \pi_\infty\| \\ & \leq \overline{\lim}_{t_0 \rightarrow \infty} \sup_{t \geq t_0} \|\pi_\infty - \pi_{t_0}\| + \overline{\lim}_{t_0 \rightarrow \infty} \sup_{t \geq t_0} \sum_{k=t_0}^{t-1} \|\pi_k - \pi_{k+1}\| \\ & \quad + \overline{\lim}_{t_0 \rightarrow \infty} \sup_{t \geq t_0} \|\pi_t - \pi_\infty\| \\ & = \overline{\lim}_{t_0 \rightarrow \infty} \sup_{t \geq t_0} \sum_{k=t_0}^{t-1} \|\pi_k - \pi_{k+1}\| \\ & = \overline{\lim}_{t_0 \rightarrow \infty} \sum_{k=t_0}^{\infty} \|\pi_k - \pi_{k+1}\| \\ & = 0. \end{aligned}$$

Because by assumption,  $\sum_{k=t_0}^{\infty} \|\pi_k - \pi_{k+1}\| < \infty$ . This completes the proof of Lemma 2.

Q.E.D.

PROOF of THEOREM 1. For any  $\eta \in \Omega$ ,

$$\begin{aligned} & \overline{\lim}_{t \rightarrow \infty} \|P(X(t) = \cdot | X(0) = \eta) - \pi_\infty\| \\ & = \overline{\lim}_{t_0 \rightarrow \infty} \overline{\lim}_{\substack{t \rightarrow \infty \\ t \geq t_0}} \left\| \sum_{\eta'} P(t, \cdot | t_0, \eta') P(t_0, \eta' | 0, \eta) - \pi_\infty \right\| \end{aligned}$$

$$\begin{aligned}
&\leq \overline{\lim}_{t_0 \rightarrow \infty} \overline{\lim}_{\substack{t \rightarrow \infty \\ t \geq t_0}} \left\| \sum_{\eta'} P(t, \cdot | t_0, \eta') P(t_0, \eta' | 0, \eta) - P(t, \cdot | t_0, \pi_\infty) \right\| \\
&\quad + \overline{\lim}_{t_0 \rightarrow \infty} \overline{\lim}_{\substack{t \rightarrow \infty \\ t \geq t_0}} \| P(t, \cdot | t_0, \pi_\infty) - \pi_\infty \|.
\end{aligned}$$

The last term is zero by Lemma 2. Furthermore, since  $P(t_0, \cdot | 0, \eta)$  and  $\pi_\infty$  has total mass 1, we have

$$\begin{aligned}
&\left\| \sum_{\eta'} P(t, \cdot | t_0, \eta') P(t_0, \eta' | 0, \eta) - P(t, \cdot | t_0, \pi_\infty) \right\| \\
&= \sum_{\omega} \left| \sum_{\eta'} P(t, \omega | t_0, \eta') P(t_0, \eta' | 0, \eta) - P(t, \omega | t_0, \pi_\infty) \right| \\
&= \sum_{\omega} \sup_{\eta''} \left| \sum_{\eta'} \{ P(t, \omega | t_0, \eta') - P(t, \omega | t_0, \eta'') \} \right. \\
&\quad \left. \{ P(t_0, \eta' | 0, \eta) - \pi_\infty(\eta') \} \right| \\
&\leq \sum_{\omega} \sup_{\eta''} \sum_{\eta'} | P(t, \omega | t_0, \eta') - P(t, \omega | t_0, \eta'') | \\
&\quad | P(t_0, \eta' | 0, \eta) - \pi_\infty(\eta') | \\
&\leq \sum_{\omega} \sup_{\eta', \eta''} | P(t, \omega | t_0, \eta') - P(t, \omega | t_0, \eta'') | \\
&\quad \sum_{\eta'} | P(t_0, \eta' | 0, \eta) - \pi_\infty(\eta') | \\
&\leq 2 \sum_{\omega} \sup_{\eta', \eta''} | P(t, \omega | t_0, \eta') - P(t, \omega | t_0, \eta'') |.
\end{aligned}$$

Finally then

$$\begin{aligned}
&\overline{\lim}_{t \rightarrow \infty} \| P(X(t) = \cdot | X(0) = \eta) - \pi_\infty \| \\
&\leq 2 \sum_{\omega} \overline{\lim}_{t_0 \rightarrow \infty} \overline{\lim}_{\substack{t \rightarrow \infty \\ t \geq t_0}} \sup_{\eta', \eta''} | P(t, \omega | t_0, \eta') - P(t, \omega | t_0, \eta'') | \\
&= 0 \quad \text{by Lemma 1.}
\end{aligned}$$

This completes the proof of Theorem 1.

Q.E.D.

We obtain the following corollaries.

COROLLARY 1. Suppose that the sampler process  $\{X(t); t = 0, 1, 2, \dots\}$  is generated by

$$P(X_s(t)=x_s, s \in S) = \pi(x_{n_t} | x_s, s \neq n_t) P(X_s(t-1)=x_s, s \neq n_t)$$

where  $\pi$  is a probability measure on  $\Omega$  such that  $0 < \pi(\omega) < 1$  for every  $\omega \in \Omega$ . Assume that for every  $s \in S$ , the sequence  $\{n_t, t \geq 1\}$  contains  $s$  infinitely often. Then for any starting configuration  $\eta \in \Omega$  and for every  $\omega \in \Omega$ ,

$$\lim_{t \rightarrow \infty} P(X(t)=\omega | X(0)=\eta) = \pi(\omega).$$

Concerning ergodicity, we use the sampler process and compute a time average of the function  $Y$ .

COROLLARY 2. Suppose that the sampler process  $\{X(t); t = 0, 1, 2, \dots\}$  is generated by

$$P(X_s(t)=x_s, s \in S) = \pi(x_{n_t} | x_s, s \neq n_t) P(X_s(t-1)=x_s, s \neq n_t)$$

where  $\pi$  is a probability measure on  $\Omega$  such that  $0 < \pi(\omega) < 1$  for every  $\omega \in \Omega$ . Assume that there exists an integer  $\tau \geq N$  such that for every  $t = 0, 1, 2, \dots$  we have

$$S \subset \{n_{t+1}, n_{t+2}, \dots, n_{t+\tau}\}.$$

Then for every function  $Y$  on  $\Omega$  and for every starting configuration  $\eta \in \Omega$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n Y(X(t)) = \int_{\omega} Y(\omega) \pi(\omega)$$

holds with probability one.

#### 4. Applications to Image Processing.

Let  $S = Z_m = \{(i,j): 1 \leq i,j \leq m\}$  denote the  $m \times m$  integer lattice; then  $F = \{F_{i,j}\}$ ,  $(i,j) \in S$ , denotes the gray levels of the original, digitized image. Lowercase letters will denote the values assumed by these random variables. It is natural to expect that the image value at a pixel does not depend on the image data outside its neighborhood, when the image data on its neighborhood is given. Specially, we model  $F$  as an MRF, or, what is the same (see Proposition 1), we assume that the probability law of  $F$  is a Gibbs distribution with respect to  $\{S, G\}$  with corresponding energy function  $U$  and potentials  $\{V_C\}$ .

Let  $H$  denote the blurring matrix corresponding to a shift-invariant point-spread function. The formulation of  $F$  gives rise to a blurred image  $H(F)$  which is recorded by a sensor. The latter often involves a nonlinear transformation of  $H(F)$ , denote here by  $\phi$ , in addition to random sensor noise  $N = \{\eta_{i,j}\}$ , which we assume to consist of independent, and for definiteness, Gaussian variables with mean  $\mu$  and standard deviation  $\sigma$ . We also assume that  $F$  and  $N$  are independent as stochastic processes. The degraded image is then a function of  $\phi(H(F))$  and  $N$ , say  $\psi(\phi(H(F)), N)$ , for example, addition or multiplication. To compute the posterior distribution, we only need to assume that  $b \rightarrow \psi(a, b)$  is invertible for each  $a$ . For notational ease, we will write

$$G = \phi(H(F)) \odot N$$

At the pixel level, for each  $(i,j) \in S$ ,

$$G_{i,j} = \phi\left(\sum_{(k,l)} H(i-k, j-l) F_{k,l}\right) \odot \eta_{i,j}.$$

Since the operation  $\odot$  is assumed invertible, we can write

$$N = \phi(G, \phi(H(F))) = \{\phi_s, s \in S\}$$

to indicate this inverse.

Let  $H_s, s \in S$ , denote the pixels which affect the blurred image  $H(F)$  at  $s$ . Observe that  $\phi_s, s \in S$ , depends only on  $g_s$  and  $\{f_t, t \in H_s\}$ . By the shift-invariance of  $H$ ,  $H_{r+s} = s + H_r$ , where  $H_r \subset S$ ,  $s+r \in S$ , and  $s + H_r$  is understood to be intersected with  $S$ , if necessary. In addition, we will assume that  $\{H_s\}$  is symmetric in that  $r+s \in H_s \Leftrightarrow -r+s \in H_s$ . Then the collection  $\{H_s \setminus \{s\}; s \in S\}$  is a neighborhood system over  $S$ . Let  $H^2$  denote the second-order system, i.e.,

$$H_s^2 = \bigcup_{r \in H_s} H_r, \quad s \in S.$$

Then it is not hard to see that  $\{H_s^2 \setminus \{s\}, s \in S\}$  is also a neighborhood system. Finally, set  $G^P = \{G_s^P, s \in S\}$  where  $G_s^P = G_s \cup H_s^2 \setminus \{s\}$ . Let  $\mu \in \mathbb{R}^N$  ( $N = m^2$ ) have all components  $= \mu$  and let  $\|\cdot\|$  denote the usual norm in  $\mathbb{R}^N$ :  $\|x\|^2 = \sum_{i=1}^N x_i^2$ .

PROPOSITION 2. For each  $g$  fixed,  $P(F=f|G=g)$  is a Gibbs distribution over  $\{S, G^P\}$  with energy function

$$U^P(f) = U(f) + \|\mu - \phi(g, \phi(H(f)))\|^2 / 2\sigma^2.$$

For a proof see [3]. The posterior distribution  $P(F=f|g)$  is a powerful tool for image analysis; in principle, we can construct the optimal estimator for the original image, examine images samples from  $P(F=f|g)$ , estimate parameters, design near-optimal statistical tests for the presence or absence of special objects, and so on. Specially, our work here is to find the value(s) of  $f$  which maximize the posterior distribution for a fixed  $g$ , i.e., minimize

$$U^P(f) = U(f) + ||\mu - \phi(g, \phi(H(f)))||^2 / 2\sigma^2, \quad f \in \Omega$$

where  $\phi$  is defined by  $\phi(H(f)) \odot \phi = g$ .

In other words, we find the value(s) of  $f$  with lowest energy.

So far, we have not discussed how these realizations from this class of Gibbs distributions are generated. Basically there are two well-known methods of generating realizations from MRF's or Gibbs distribution's. They are the "exchange" type and the "spin-flip" type algorithms. In the exchange type algorithm, also known as the Metropolis' Algorithm [5], two pixels are chosen at random. Their values are exchanged if they are different and if the exchange will take the system to a more probable (lower energy) configuration. If the new configuration is less probable then the exchange will or will not take place depending on the comparison of the ratio of the probabilities of the new and the old configurations with a random number uniform on  $[0,1]$ . The randomization is necessary to ensure that the system does not get stuck in a local high probability configuration. The ratio of the probabilities of the new and the old configurations are calculated easily due to the Gibbs distribution formulation, without actually determining either of the probabilities, which would be extremely difficult.

It is well known that this algorithm will converge to a configuration that maximizes the joint probability but the rate of convergence is a difficult problem of statistical physics and a completely satisfactory solution to this problem does not exist. As would be expected, the initial configuration does not influence the convergence properties of the algorithm. It might only take a few more iterations to converge for certain initial configurations as compared to others. The

Metropolis' Algorithm has been used by Cross and Jain [2] in generating textures using MRF models. In a recent paper, Kirkpatrick, Gelatt and Vecchi [4] proved a stochastic approximation method for solving combinatorial optimization problems, which can be used for the minimization problem. A major drawback of the Metropolis' Algorithm is that the number of pixels in each gray level does not change during iterations of the algorithm, due to the fact that new configurations are generated simply by exchanging two pixel values. Therefore we choose to use the second method for generating realizations from a MRF (or Gibbs distribution).

The second set of algorithms for generating realizations from a MRF (or Gibbs distribution) are known as "spin-flip" algorithms. Recently a version of this algorithm is presented by Geman and Geman [3] in an image processing context. This algorithm, which they called the "Gibbs Sampler", works as follows. A pixel is chosen at random or in a deterministic manner. The value of the pixel is renewed by disregarding its present value, noting the values in its neighborhood and replacing the pixel value by a random number generated according to the conditional distribution specified by the MRF (or Gibbs distribution). Pixel visiting mechanism can be random or deterministic, such as raster scan, the important point being that each pixel should be visited infinitely often as the algorithm proceeds ad infinitum. This algorithm has the characteristics of a relaxation algorithm in image processing; therefore, it is also called stochastic relaxation.

In this section we use the sampler process stated in §2 to find the value(s) of  $f \in \Omega$  with lowest energy.

Let us indicate the dependence of  $\pi$  on  $T$  by writing  $\pi_T$ , and let  $T(t)$  denote the temperature at stage  $t$ .

We assume that  $\pi_{T(t)}$  is Gibbs distribution. Let

$$\Omega_0 = \{\omega \in \Omega: U(\omega) = \min_{\eta} U(\eta)\}$$

and let  $\pi_0$  be the uniform distribution on  $\Omega_0$ .

Finally define

$$U^* = \max_{\omega} U(\omega),$$

$$U_* = \min_{\omega} U(\omega),$$

$$\Delta = U^* - U_*$$

The sampler process  $\{X(t); t = 0, 1, 2, \dots\}$  is generated by

$$P(X_S(t)=x_S, s \in S) = \pi_{T(t)}(x_{n_t} | x_S, s \neq n_t) P(X(t-1)=x_S, s \neq n_t).$$

We obtain the following

**THEOREM 2.** Assume that there exists an integer  $\tau \geq N$  such that for every  $t = 0, 1, 2, \dots$  we have

$$S \subset \{n_{t+1}, n_{t+2}, \dots, n_{t+\tau}\}.$$

Let  $T(t)$  be any decreasing sequence of temperatures for which

$$a) \quad T(t) \rightarrow 0 \text{ as } t \rightarrow \infty$$

$$b) \quad T(t) \geq N\Delta/\log t \text{ for all } t \geq t_0 \text{ for some integer } t_0 \geq 2.$$

Then for any starting configuration  $\eta \in \Omega$  and for every  $\omega \in \Omega$ ,

$$\lim_{t \rightarrow \infty} P(X(t)=\omega | X(0)=\eta) = \pi_0(\omega).$$

**OUTLINE OF PROOF.** We replace  $\pi_t$  and  $\pi_\infty$  in Theorem 1 for  $\pi_{T(t)}$  and  $\pi_0$  in Theorem 2. By assumption, we obtain the inequality

$$\delta(t) \geq \frac{1}{L} \exp\left(-\frac{\Delta}{T(t)}\right) \geq \frac{1}{L} \exp\left(-\frac{\log t}{N}\right) = \frac{1}{L} \frac{1}{N^{\frac{1}{\sqrt{t}}}}.$$

It is easy to show the convergence of  $\pi_{T(t)}$  to  $\pi_0$ . Because

$$\begin{aligned}\pi_{T(t)}(\omega) &= \frac{\exp(-\frac{U(\omega)}{T(t)})}{\sum_{\omega' \in \Omega_0} \exp(-\frac{U(\omega')}{T(t)}) + \sum_{\omega' \in \Omega \setminus \Omega_0} \exp(-\frac{U(\omega')}{T(t)})} \\ &= \frac{\exp(-\frac{U(\omega) - U_*}{T(t)})}{|\Omega_0| + \sum_{\omega' \in \Omega \setminus \Omega_0} \exp(-\frac{U(\omega') - U_*}{T(t)})}\end{aligned}$$

converges to  $\pi_0$  as  $t \rightarrow \infty$ .

We obtain Theorem 2 as the result corresponding to Theorem 1.

Q.E.D.

## References

- [1] J. Besag, Spacial interaction and the statistical analysis of lattice systems, J. Royal Stat. Soc., B 34(1972), 75-83.
- [2] G. C. Cross and A. K. Jain, Markov random field texture models, IEEE Trans. PAMI, 5(1983), 25-39.
- [3] S. Geman and D. Geman, Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images, IEEE Trans. PAMI, 6(1984), 721-741.
- [4] S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi, Optimization by simulated annealing, Science, 220(1983), 671-680.
- [5] N. Metropolis et. al., Equation of state calculations by fast computing mechanics, J. Phys. Chem., 21(1953), 1087-1092.
- [6] F. Spitzer, Markov random fields and Gibbs ensembles, Amer. Math. Monthly, 78(1971), 142-154.

END

DTIC

9-86